Évolutions de la pratique statistique dans l'enseignement du second degré en France :

du calcul statistique au développement de la pensée statistique

Philippe DUTARTE - Congrès CFIES Bruxelles septembre 2010

1 - La statistique dans les programmes d'enseignement français depuis une vingtaine d'années

Les programmes des sections de techniciens supérieurs (1988)

- STATISTIQUE INFÉRENTIELLE

  « Sous l'impulsion notamment du **mouvement de la qualité**, les méthodes statistiques sont aujourd'hui largement utilisées dans les milieux économique, social ou professionnel.

  Il s'agit de faire percevoir, à partir d'exemples figurant au programme, ce que sont les procédures de décision en univers aléatoire, ainsi que leur pertinence. »
- FIABILITÉ
- PLANS D'EXPÉRIENCE

Les programmes de seconde de 2000

- INTRODUCTION

  « L'**informatique**, devenue aujourd'hui absolument incontournable, permet de rechercher et d'observer des lois expérimentales… […] Cette possibilité d'**expérimenter**, classiquement davantage réservée aux autres disciplines, doit ouvrir largement la dialectique entre l'observation et la démonstration, et, sans doute à terme, changer profondément la nature de l'enseignement. »

  « Un des apports majeurs de l'informatique réside aussi dans la puissance de **simulation** des ordinateurs ; la simulation est ainsi devenue une pratique scientifique majeure : une approche en est proposée dans le chapitre statistique. »

Les programmes de seconde de 2000

- STATISTIQUE

  « En seconde le travail sera centré sur :

  – la **réflexion** conduisant au choix de résumés numériques d'une série statistique quantitative ;

  – la notion de **fluctuation d'échantillonnage** vue ici sous l'aspect élémentaire de la variabilité de la distribution des fréquences ;

  – la **simulation** à l'aide du générateur aléatoire d'une calculatrice. »

  « L'enseignant traitera des **données en nombre suffisant** pour que cela justifie une étude statistique ; il proposera des sujets d'étude et des simulations en fonction de l'intérêt des élèves, de l'actualité et de ses goûts. »

  « L'objectif est de faire **réfléchir** les élèves sur la nature des données traitées, et de s'appuyer sur des représentations graphiques pour **justifier des choix** de résumé. »

Le programme de première scientifique (2001)

- « Le programme de la classe de première introduit quelques outils descriptifs nouveaux […]. Ces éléments de statistique pourront notamment être travaillés pour des séries construites à partir de **séries simulées**; on rencontre ainsi des répartitions variées et **on prépare la notion d'estimateur**. **Cette partie descriptive ne doit pas faire l'objet de longs développements numériques, ni être déconnectée du reste du programme de probabilité et statistique.** »
- « Le **lien entre loi de probabilité et distributions des fréquences** sera éclairé par un énoncé vulgarisé de la loi des grands nombres. »

Le programme de terminale scientifique (2002)

- « À titre indicatif, la **répartition horaire** des différents chapitres peut être : analyse 45 % (environ 14 semaines), géométrie 35 % (environ 11 semaines), probabilités et statistique **20 %** (environ 6 semaines). »
- « On vise aussi, en complément à l'usage des simulations introduit dès la seconde, une première sensibilisation à d'autres classes de problèmes, notamment celui de l'**adéquation d'une loi de probabilité à des données expérimentales**. »

  « L'élève devra être capable de poser le problème de l'adéquation à une loi équirépartie et de se reporter à des résultats de simulation qu'on lui fournit. »

Les programmes des sections techniques tertiaires et socio-médicales de première et terminale (2007-2008)

- « On privilégiera les données issues des autres disciplines ou des médias, afin de développer une **attitude critique** vis-à-vis des informations chiffrées. L'emploi du **tableur** et de la calculatrice facilite l'exploitation des données. »
- Le programme de statistique fournit un terrain pour des **activités pluridisciplinaires** et pour la consolidation des techniques élémentaires de calcul : pourcentages, proportionnalité, usage de fractions…
- L'étude de ce chapitre doit constituer **un moment important de la formation** des élèves (développement de l'esprit critique, capacité à analyser les résultats d'une enquête…). Il est donc nécessaire que les élèves disposent d'un temps suffisant pour se familiariser avec ces notions.

Les programmes des sections techniques tertiaires et socio-médicales de première et terminale (2007-2008)

- « Le tableur permet d'aborder des populations de grand effectif. »
- « Le calcul d'un écart-type **à la main** à l'aide d'un tableau n'est pas un objectif et ne peut être exigible. »
- « Rédiger l'**interprétation d'un résultat** ou l'analyse d'un graphique : cette capacité est attendue pour l'intégralité de l'étude des séries de données statistiques à une variable. »

L'introduction de l'aléatoire dans les programme de collège (2008)

- « La notion de probabilité est **abordée à partir d'expérimentations** qui permettent d'**observer les fréquences** des issues dans des situations familières (pièces de monnaie, dés, roues de loteries, urnes, etc.).
- La notion de probabilité est utilisée pour modéliser des situations simples de la vie courante. Les situations étudiées concernent les expériences aléatoires à une ou *à deux épreuves.* »

Le programme de seconde professionnelle (2009)

- « Ce domaine [la statistique] constitue un **enjeu essentiel de formation du citoyen**. Il s'agit de fournir des outils pour comprendre le monde, décider et agir dans la vie quotidienne. »
- « La notion de fluctuation d'échantillonnage, essentielle en statistique, est abordée dans cette partie du programme en étudiant la variabilité d'observation d'une fréquence. Elle favorise une expérimentation de l'aléatoire. L'objectif de ce module est de faire comprendre que le hasard suit des lois et de préciser l'approche par les fréquences de la notion de probabilité initiée en classe de troisième. »

Le programme de première professionnelle (2010)

« La connaissance de la « variabilité naturelle » des fréquences d'échantillons (la probabilité qu'un échantillon aléatoire de taille *n* fournisse une fréquence dans l'intervalle $\left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}} \right]$ est supérieure à 0,95) permet de juger de la pertinence de certaines observations. »

Le programme de seconde (2009 - 2010)

- ANALYSE DE DONNÉES

  « L'objectif est de faire réfléchir les élèves sur des **données réelles, riches et variées** (issues, par exemple, d'un fichier mis à disposition par l'INSEE), synthétiser l'information et **proposer des représentations pertinentes**. »

- ÉCHANTILLONNAGE

  « L'objectif est d'amener les élèves à un questionnement lors des activités suivantes :

  – l'**estimation d'une proportion inconnue** à partir d'un échantillon ;

  – la **prise de décision à partir d'un échantillon**. »

Les projets de programme de première générale et technologique industrielle (2011)

- CONTENU

  « Utilisation de la loi binomiale pour une prise de décision à partir d'une fréquence. »
- CAPACITES ATTENDUES

  « Exploiter un **intervalle de fluctuation déterminé à l'aide de la loi binomiale** pour rejeter ou non une **hypothèse sur une proportion**. »
- COMMENTAIRES

  « L'objectif est d'amener les élèves à expérimenter la notion de « différence significative » par rapport à une valeur attendue et à remarquer que pour une taille de l'échantillon importante, **on conforte les résultats vus en classe de seconde**. »

2 – Analyse des manuels de seconde générale proposés à la rentrée 2010

(chapitre « Statistique descriptive – Analyse des données »)

Exercices calculatoires

(sans utilisation des calculs ou graphiques effectués)

Editeur « Collection »	
Nathan « Hyperbole »	12/52 soit 23 %
Nathan « Transmath »	9/57 soit 16 %
Nathan « Antibi »	45/59 soit 76 %
Bordas « Indice »	20/25 soit 80 %
Bordas « Pixel »	29/58 soit 50 %
Didier « Math'x »	18/82 soit 22 %
Hachette « Déclic »	28/64 soit 44 %
Hachette « Repères »	32/55 soit 58 %
Hatier « Odyssée »	12/44 soit 30 %

Nathan « Antibi » page 237

27 * On donne la série statistique suivante :

Valeurs	1	a	5	$2a$	9
Effectifs	28	22	14	20	16

a) Déterminez a sachant que la moyenne est égale à 4,59.

b) Calculez la médiane de cette série.

Le « cours » comporte-t-il autre chose que des définitions et des formules de calcul ?

Editeur « Collection »	
Nathan « Hyperbole »	Oui
Nathan « Transmath »	Non
Nathan « Antibi »	Non
Bordas « Indice »	Non
Bordas « Pixel »	Oui
Didier « Math'x »	Oui
Hachette « Déclic »	Oui
Hachette « Repères »	Oui
Hatier « Odyssée »	Non

Bordas « Indice » pages 112-113

COURS

Le rôle de la Statistique descriptive est de présenter une masse de données sous forme lisible. Puis si possible, de la résumer par quelques nombres caractéristiques (moyenne, médiane…).

1 Le vocabulaire

Population. Une population est un ensemble de personnes ou d'objets, appelés individus, définis par une propriété commune.

Par exemple, les habitants d'un pays, les automobiles fabriquées en 2000.

Caractère. Pour une population choisie, on peut étudier un caractère de ses individus.

Par exemple, on peut étudier le caractère « taille » des élèves d'un lycée.

Caractère quantitatif. Un caractère est dit quantitatif lorsqu'il est possible de le mesurer en associant un nombre à chaque individu. Un caractère quantitatif est aussi appelé **variable.**

Par exemple, l'âge, la taille, le nombre de frères et sœurs.

– Un caractère quantitatif est dit **continu** lorsque les nombres qui le mesurent peuvent prendre, a priori, toutes les valeurs d'un intervalle.

Par exemple, le poids, la taille, la durée de vie d'un moteur.

– Il est dit **discret** dans le cas contraire.

Par exemple, l'année de naissance, le nombre d'enfants par famille…

Caractère qualitatif. On appelle ainsi tout caractère non quantitatif.

Par exemple, la couleur des yeux.

Dans ce chapitre, les séries étudiées sont des séries dont le caractère est quantitatif.

2 Présentation d'une série statistique

2.1 Effectifs. Effectifs cumulés croissants

• Vous avez déjà rencontré la présentation d'une série statistique par un tableau d'effectifs. Le tableau 1 a été dressé après recensement du nombre d'enfants dans chacun des foyers d'un village.

Tableau 1

Nombre d'enfants x_i	0	1	2	3	4	5	6	7
Effectif n_i	290	170	155	95	43	27	20	10

Si on note N l'effectif total, alors $n_1 + n_2 + \ldots + n_8 = N = 810$.

• À partir de ce tableau, on peut dresser le **tableau des effectifs cumulés croissants** (en abrégé : e.c.c.).

Tableau 2

Nombre d'enfants $x \leq$	0	1	2	3	4	5	6	7
e.c.c.	290	460	615	710	753	780	800	810

Dans le tableau 1, on lit que 155 foyers ont 2 enfants. Dans le tableau 2, on lit que dans 615 foyers, le nombre d'enfants est inférieur ou égal à 2. (Et $615 = n_1 + n_2 + n_3 = 290 + 170 + 155$.)

Intérêt du tableau des effectifs cumulés croissants. Ce tableau donne directement le nombre de foyers qui ont au plus x_i enfants.

Par exemple, pour $x_i = 2$, il y a 615 foyers qui ont au plus 2 enfants.

COURS

2.2 Fréquences. Fréquences cumulées croissantes

• À partir des effectifs, on peut dresser le tableau des fréquences.

Par définition, $f_i = \frac{n_i}{N}$. Ainsi, $f_1 = \frac{n_1}{N} = \frac{290}{810} \approx 0,36$.

x_i	0	1	2	…
f_i	0,36	0,21	0,19	…

La **somme des fréquences** est toujours égale à 1.

• Puis à partir du tableau des fréquences, on peut dresser le **tableau des fréquences cumulées croissantes** (f.c.c.).

$x \leq$	0	1	2	
f.c.c.	0,36	0,57	0,76	…

3 Paramètres d'une série statistique

3.1 Moyenne

Définition 1 La **moyenne** d'une série statistique est le nombre, noté $\bar{x}$, défini par :

$$\bar{x} = \frac{n_1x_1 + n_2x_2 + \ldots + n_px_p}{N}.$$

Variable x_i	x_1	x_2	…	x_p
Effectif n_i	n_1	n_2	…	n_p

Effectif total : $N = n_1 + n_2 + \ldots + n_p$

On peut aussi calculer $\bar{x}$ en utilisant les fréquences f_i : $\bar{x} = f_1x_1 + f_2x_2 + \ldots + f_px_p$.

En effet, $\bar{x} = \frac{n_1x_1 + n_2x_2 + \ldots + n_px_p}{N}$, et puisque $f_i = \frac{n_i}{N}$, on obtient : $\bar{x} = f_1x_1 + f_2x_2 + \ldots + f_px_p$.

3.2 Médiane, quartiles

Définition 2 La **médiane** d'une série statistique est le nombre, noté Me, tel que :

• 50 % au moins des individus ont une valeur du caractère inférieure ou égale à Me ;

• 50 % au moins des individus ont une valeur supérieure ou égale à Me.

Pratiquement. La liste des N valeurs étant rangée **par ordre croissant**, chacune figurant un nombre de fois égal à son effectif :

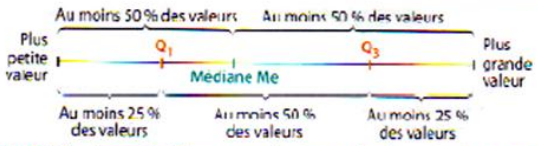
– si N est impair, $N = 2k + 1$, alors Me est la valeur située au rang $k + 1$;

– si N est pair, $N = 2k$, alors Me est la demi-somme des valeurs de rang k et de rang $k + 1$.

Définition 3 La liste des N valeurs est rangée par **ordre croissant.**

• **Le premier quartile** est la plus petite valeur Q_1 de la liste telle qu'au moins un quart des valeurs de la liste sont **inférieures ou égales** à Q_1.

• **Le troisième quartile** est la plus petite valeur Q_3 de la liste telle qu'au moins les trois quarts des valeurs de la liste sont **inférieures ou égales** à Q_3.


Au moins 50 % des valeurs
Au moins 50 % des valeurs
Plus petite valeur
Q_1
Q_3
Plus grande valeur
Médiane Me
Au moins 25 % des valeurs
Au moins 50 % des valeurs
Au moins 25 % des valeurs

Remarque. L'**étendue** est la différence entre la plus grande valeur et la plus petite valeur.
L'**écart interquartile** est le nombre $Q_3 - Q_1$.

Exercices traitant des données « réelles »

Editeur « Collection »	
Nathan « Hyperbole »	15/52 soit 29 %
Nathan « Transmath »	23/57 soit 40 %
Nathan « Antibi »	12/59 soit 20 %
Bordas « Indice »	3/25 soit 12 %
Bordas « Pixel »	4/58 soit 7 %
Didier « Math'x »	24/82 soit 29 %
Hachette « Déclic »	5/64 soit 8 %
Hachette « Repères »	8/55 soit 15 %
Hatier « Odyssée »	13/44 soit 30 %

Bordas « Pixel » page 245

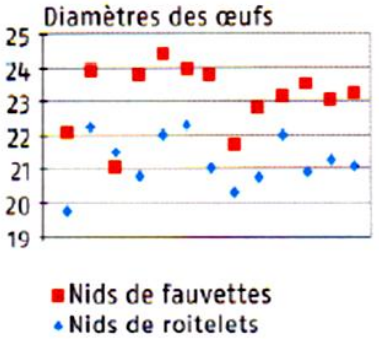
Un oiseau parasite

Le coucou est un oiseau qui ne construit pas de nid et pond ses œufs dans des nids déjà existants. Une étude a été faite sur des œufs de coucous déposés dans des nids de roitelets (de petite taille) et des nids de fauvette (de grande taille). Pour chaque œuf de coucou prélevé, on a mesuré le diamètre en millimètres.

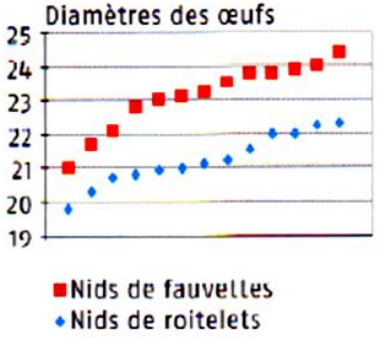
Nids de roitelets	19,8	22,2	21,5	20,8	22	22,3	21	20,3	20,7	22	20,9	21,2	21,1
Nids de fauvettes	22,1	23,9	21	23,8	24,4	24	23,8	21,7	22,8	23,1	23,5	23	23,2

1. Les données brutes suggèrent-elles des remarques?

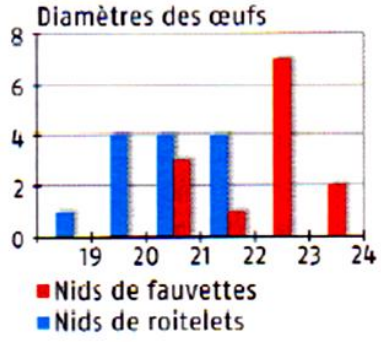
Pour ces données, les graphiques suivants ont été obtenus avec un tableur:



Graphique 1



Graphique 2



Graphique 3

2. Comment chaque graphique a-t-il été construit?

3. Au vu des graphiques, émettre une hypothèse concernant les œufs de coucou. Certains graphiques sont-ils plus parlants que d'autres?

4. Pour chaque série, calculer la moyenne, la médiane, l'étendue; les résultats confortent-ils l'hypothèse émise?

Nathan « Hyperbole » page 238

SUJETS D'EXPOSÉS

SUJET 1

Adolphe Quételet, mathématicien et astronome du XIX[e] siècle, utilisait les statistiques pour définir un « homme moyen ».

→ Rechercher sur Internet des informations concernant l'homme moyen de Quételet et discuter de la possibilité qu'il existe.

Lors d'un exposé, présenter vos recherches à la classe.

SUJET 2

MPS Méthodes et pratiques scientifiques

Rechercher sur Internet des données sur l'évolution, durant les trente dernières années, de la production de déchets par habitant en France.

→ Après avoir analysé ces données (type de déchets, durée de vie dans le sol, pourcentage de recyclage…), utiliser les statistiques descriptives pour présenter vos résultats à la classe.

Activités d'analyse d'un grand nombre de données

Editeur « Collection »	
Nathan « Hyperbole »	3 (adresses Internet)
Nathan « Transmath »	2 (adresses Internet)
Nathan « Antibi »	1 (« recopie d'un tableau »)
Bordas « Indice »	0
Bordas « Pixel »	1 (adresse Internet)
Didier « Math'x »	6 (site associé)
Hachette « Déclic »	0
Hachette « Repères »	0
Hatier « Odyssée »	2

Nathan « Hyperbole » page 232

25 B2i L 3-4 L 3-5 **a)** Sur internet, saisir l'adresse :

http://www.insee.fr/fr/ffc/figure/NATTEF06207.xls

et ouvrir le fichier « Insécurité routière ».

Dans cet exercice, on ne s'intéresse qu'aux « Victimes blessées ».

b) Quelle formule faut-il saisir en cellule Y18 pour obtenir le nombre moyen des victimes blessées sur ces vingt-trois années ?

Étirer cette formule jusqu'en Y25.

c) Déterminer les fréquences correspondant à chaque type de blessés pour les années 1985 et 2007.

d) À l'aide de l'assistant graphique, représenter les résultats du **c)** à l'aide d'un diagramme en barres.

Commenter la comparaison de ces deux années.

Didier « Math'x » page 190

90 TICE **Les chiffres du chômage**

Le fichier disponible sur le site fournit le nombre trimestriel de chômeurs en France de 1975 à 2008.

1. a. Choisir un type de graphique permettant d'illustrer l'évolution du chômage, puis le réaliser.

b. Commenter le graphique obtenu.

2. Calculer des résumés numériques adaptés.

3. Le nombre de chômeurs au 1er trimestre 2009 est environ 2 455 000. En utilisant les indicateurs précédents, situer cette valeur par rapport à celles de 1975 à 2008.

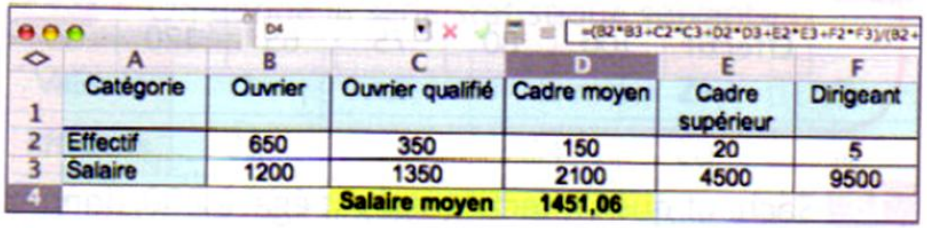
Activités d'expérimentation avec ordinateur

Editeur « Collection »	
Nathan « Hyperbole »	5 (dont 1 TP)
Nathan « Transmath »	4
Nathan « Antibi »	3
Bordas « Indice »	0 (2 TP « mode d'emploi »)
Bordas « Pixel »	1
Didier « Math'x »	8 (dont 3 TP)
Hachette « Déclic »	1
Hachette « Repères »	0 (1 exo « mode d'emploi »)
Hatier « Odyssée »	3 (dont 2 TP)

Nathan « Antibi » page 235

8 * **À l'aide d'un tableur**

On a représenté sur un tableur les effectifs (ligne 2) et les salaires (ligne 3) des employés d'une PME, en fonction de leur catégorie.



	A	B	C	D	E	F
1	Catégorie	Ouvrier	Ouvrier qualifié	Cadre moyen	Cadre supérieur	Dirigeant
2	Effectif	650	350	150	20	5
3	Salaire	1200	1350	2100	4500	9500
4			Salaire moyen	1451,06		

a) La formule qui permet de calculer le salaire moyen a été coupée. Complétez cette formule.

b) Reproduisez le tableau ci-dessus sur un tableur, en plaçant en E4 la formule complétée au **a)**.
En validant, vous devez obtenir 1 451,06.

c) Montrez, en effectuant plusieurs essais, qu'en diminuant tous les salaires (ligne 3) et en modifiant convenablement les effectifs (ligne 2), il est possible d'augmenter le salaire moyen.

Note
Le tableur permet d'obtenir instantanément le salaire moyen après chaque modification de la configuration.

Didier « Math'x » page 176

Travaux pratiques

3 En autonomie

TICE @ @

OBJECTIF Proposer des résumés graphiques et numériques pertinents.

Ouvrir le fichier disponible sur le site fournissant le nombre de nuitées dans l'hôtellerie, par département de France métropolitaine (*Source* : Direction du tourisme).

1. Représenter l'évolution du nombre total de nuitées dans l'hôtellerie en France métropolitaine de 1997 à 2007 et analyser le graphique obtenu.
Appeler le professeur pour exposer vos résultats.

2. Comparer les deux départements (hors Paris) dont la dispersion des valeurs est la plus grande et la plus faible. Utiliser des résumés graphiques et numériques pertinents.
Appeler le professeur pour exposer votre démarche et vos résultats.

3. Donner une représentation graphique montrant la répartition, par département, des nuitées en hôtel en 2007. Choisir l'indicateur de tendance central adapté à la situation.
Appeler le professeur pour exposer votre démarche et vos résultats.

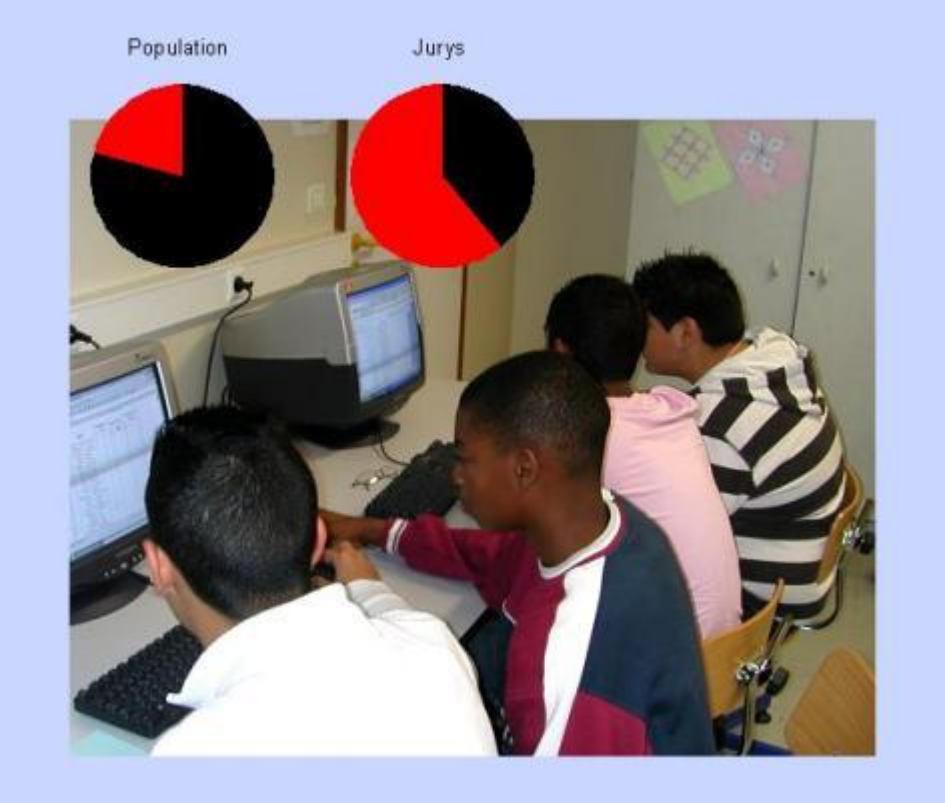
3 - Le rôle de l’outil informatique et de la démarche d’investigation

Simulation

- Approche fréquentiste de la notion de probabilité.
- Compréhension de la fluctuation d'échantillonnage.

Exemple :

l'affaire « Castaneda »



Énoncé distribué aux élèves

L'affaire Castaneda contre Partida.

En Novembre 1976 dans un comté du sud du Texas, Rodrigo Partida était condamné à huit ans de prison.

Il attaqua ce jugement au motif que la désignation des jurés de ce comté était discriminante à l'égard des Américains d'origine mexicaine. Alors que **79,1%** de la population du comté était d'origine mexicaine, sur les **870** personnes convoqués pour être jurés lors des 11 années précédentes, il n'y eût que **339** personnes d'origine mexicaine.

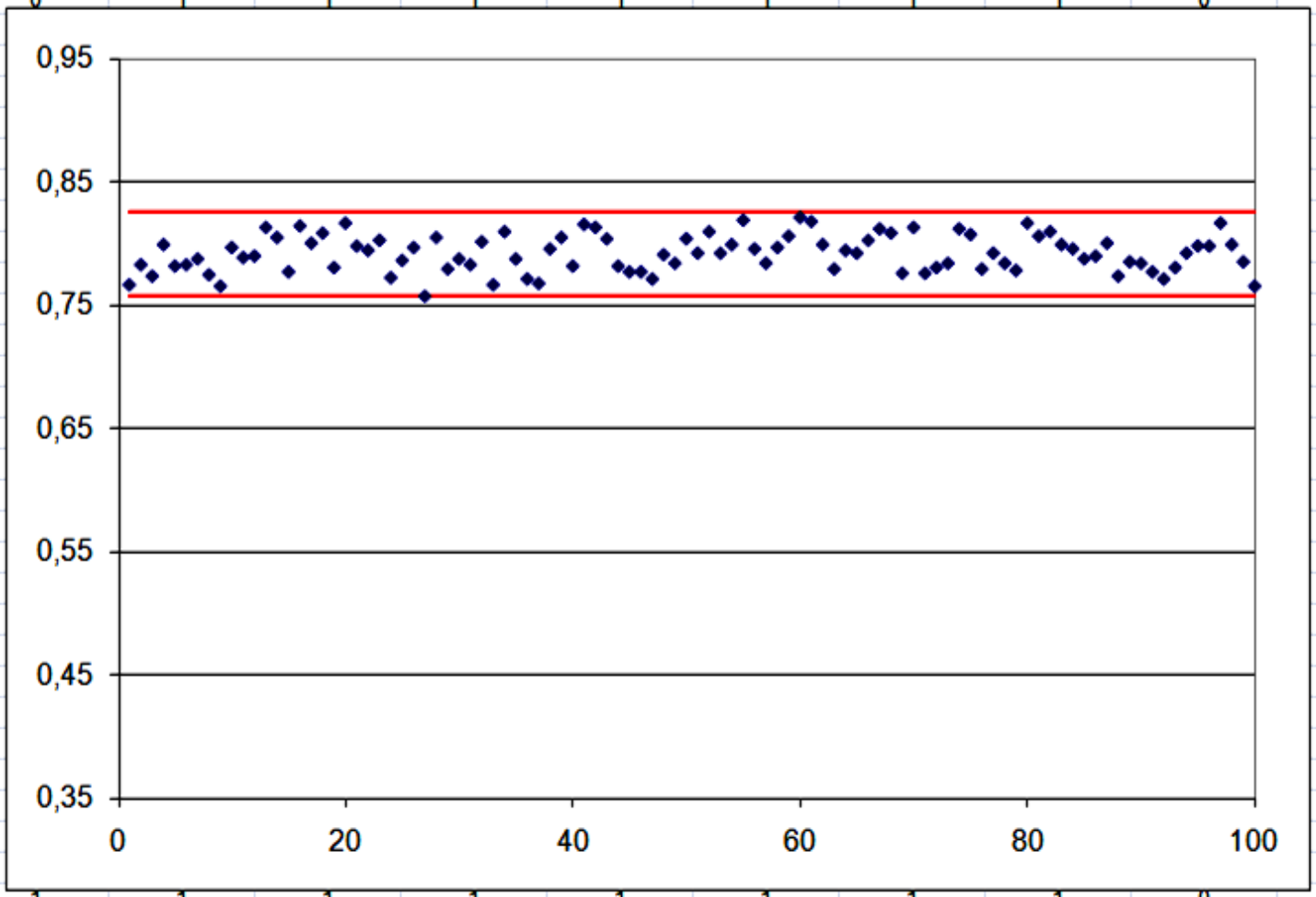
Produisez votre expertise statistique.

Devant la Cour Suprême, un expert statisticien produisit des arguments pour convaincre du bien fondé de la requête de l'accusé. En vous situant dans le rôle de cet expert, produisez à votre tour des calculs, des raisonnements, des graphiques... pour montrer que le hasard ne peut pas « raisonnablement » expliquer à lui seul la sous-représentation des américains d'origine mexicaine dans les jurys de ce comté.

- Vous commencez ce travail en binômes en utilisant les documents disponibles, la calculatrice, le tableur.
- Vous terminez la rédaction (arguments en français, calculs, graphiques...) en devoir individuel, à la maison.

A843 | =ENT(ALEA()+0,791)

	A	B	C	D	E	F	G	H	I	J	K
839	1	1	1	1	0	1	1	0	1	1	1
840	0	0	1	0	1	1	0	1	0	1	1
841	1	1	1	1	1	1	1	1	1	1	0
842	1	0	1	1	1	1	1	1	1	0	0
843	0										1
844	1										1
845	1										1
846	1										0
847	1										1
848	1										1
849	1										1
850	0										0
851	1										0
852	1										1
853	1										1
854	1										1
855	1										1
856	1										1
857	1										1
858	1										1
859	1										1
860	1										0
861	1										1
862	0										1
863	1										1
864	1										0
865	1										1
866	1										1
867	1										1
868	0										0
869	1										1
870	1										1
871	1	1	1	1	1	1	1	1	1	0	1
872	**0,77**	**0,78**	**0,77**	**0,80**	**0,78**	**0,78**	**0,79**	**0,77**	**0,77**	**0,80**	**0,79**
873											
874	**Observation dans les jurys : 0,39**										



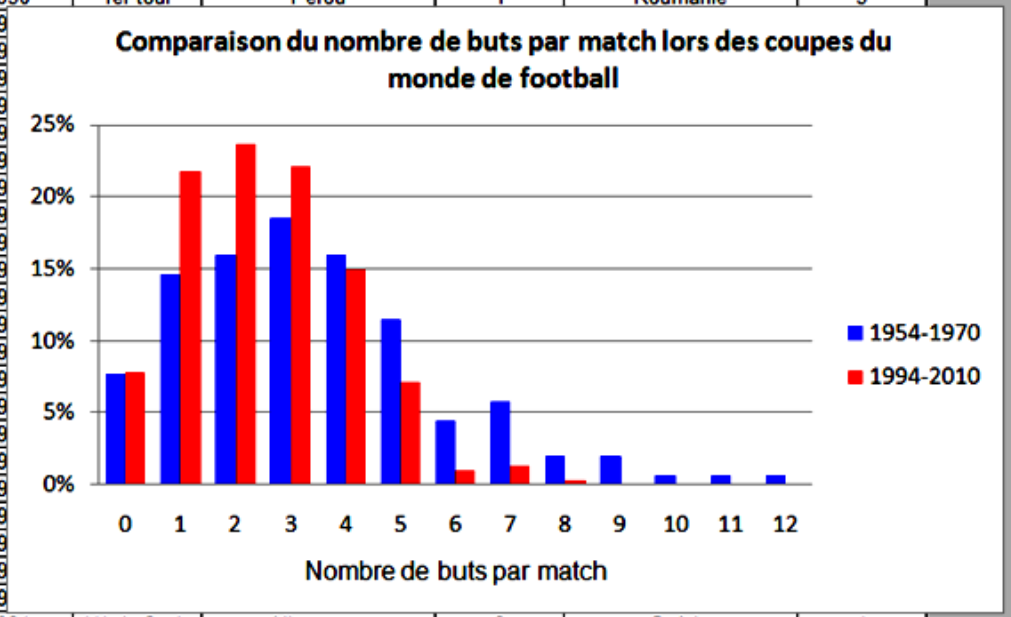
Accès à des données « réelles », Facilités de tri, calcul, illustration

Exemple : les scores des matchs des coupes du Monde de football.

L14 | =NB.SI(I$80:I$236;"="&K14)/157

Buts marqués durant les coupes du monde de football de 1930 à 2010

Année	Niveau de la compétition	Pays	Buts	Pays	Buts	Tirs au but	Nb de buts par match (hors tirs au but)
1930	1er tour	France	4	Mexique	1		5
1930	1er tour	Belgique	0	Usa	3		3
1930	1er tour	Brésil	1	Yougoslavie	2		3
1930	1er tour	Pérou	1	Roumanie	3		4
19							1
19							3
19							4
19							3
19							1
19							1
19							9
19							4
19							1
19							4
19							4
19							7
19							7
19							6
19							8
19							3
19							4
19							5
19							5
19							7
19							5
19							6
1934	1/4 de finale	Allemagne	2	Suède	1		3
1934	1/4 de finale	Suisse	2	Tchécoslovaquie	3		5
1934	1/4 de finale	Autriche	2	Hongrie	1		3
1934	1/4 de finale	Espagne	1	Italie	1		2



Comparaison 1954-1970 et 1994-2010

Nb de buts	1954-1970	1994-2010
0	7,6%	7,8%
1	14,6%	21,8%
2	15,9%	23,7%
3	18,5%	22,1%
4	15,9%	14,9%
5	11,5%	7,1%
6	4,5%	1,0%
7	5,7%	1,3%
8	1,9%	0,3%
9	1,9%	0,0%
10	0,6%	0,0%
11	0,6%	0,0%
12	0,6%	0,0%

	1954-1970	1994-2010
moyenne	3,43	2,48
écart type	2,36285988	1,53415533
Q1	2,00	1,00
Mé	3,00	2,00
Q3	5,00	3,00

Expérimentation numérique

Exemple 1 : algorithme de recherche d'un intervalle de fluctuation en utilisant la loi binomiale.

C6 : =SI(B6>0,025;SI(B6>0,975;"";A6/B$3);"")

Intervalle de fluctuation à 95 % d'une fréquence donné par la loi binomiale

n = taille de l'échantillon p = proportion supposée dans la population

n= 50 **p= 0,3** p-1/rac(n)= 0,158579 p+1/rac(n)= 0,441421

k	P(X<=k)	Intervalle de fluctuation à environ 95 %
0	1,79847E-08	
1	4,0337E-07	
2	4,44992E-06	
3	3,21977E-05	
4	0,000171927	
5	0,000722862	
6	0,002493722	
7	0,007264203	
8	0,018253347	
9	0,040231634	0,18
10	0,078850625	0,2
11	0,139036065	0,22
12	0,222865785	0,24
13	0,327883236	0,26
14	0,446831574	0,28
15	0,569178436	0,3
16	0,683878619	0,32
17	0,782193062	0,34
18	0,859440124	0,36
19	0,915197401	0,38
20	0,952236165	0,4
21	0,974912958	0,42
22	0,987723874	

On peut modifier le contenu des cellules B3 (valeur de n, ici inférieure ou égale à 1000) et D3 (valeur de p).
"L'algorithme" est donné par la fonction SI introduite en C6 puis recopiée vers le bas.

Pour n supérieur à 30, n*p et n*(1-p) supérieurs à 5, les résultats sont comparables à ceux admis en seconde.
(La majoration de l'intervalle donnée en seconde est rappelée en F3 et H3).

Exemple 2 : seuil de signification.

On mesure en laboratoire avec un ***compteur Geiger*** un objet pouvant être « radioactif ». Le compteur est réglé selon une certaine sensibilité et on effectue une mesure à 1 m de l'objet, pendant 10 s. L'instrument compte 37 désintégrations ou « coups ». Cependant, avec ce réglage et dans ces conditions, une mesure de « ***bruit de fond*** » (correspondant à l'environnement du laboratoire) donne en moyenne un comptage de 30 coups. La question qui se pose est de savoir si la différence observée est assez importante pour considérer l'objet comme « radioactif ».

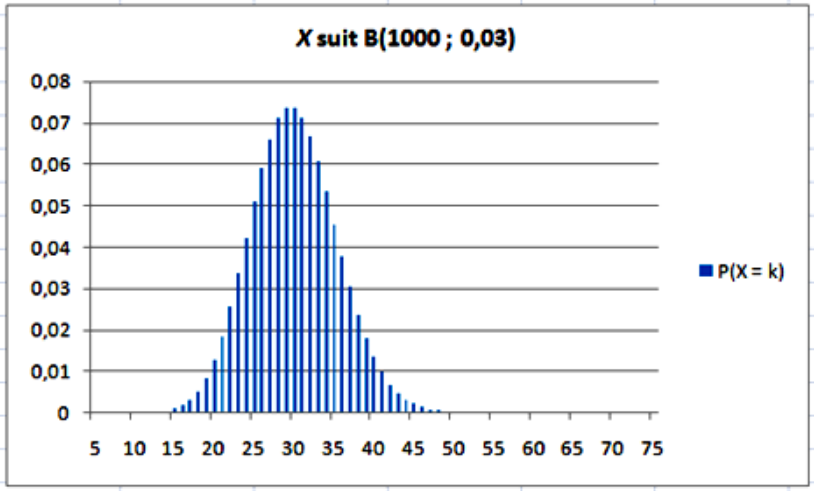
On peut simuler le nombre de désintégrations pendant une durée donnée (correspondant théoriquement à une loi de Poisson) à l'aide d'une loi binomiale, et ainsi expérimenter la notion de seuil.

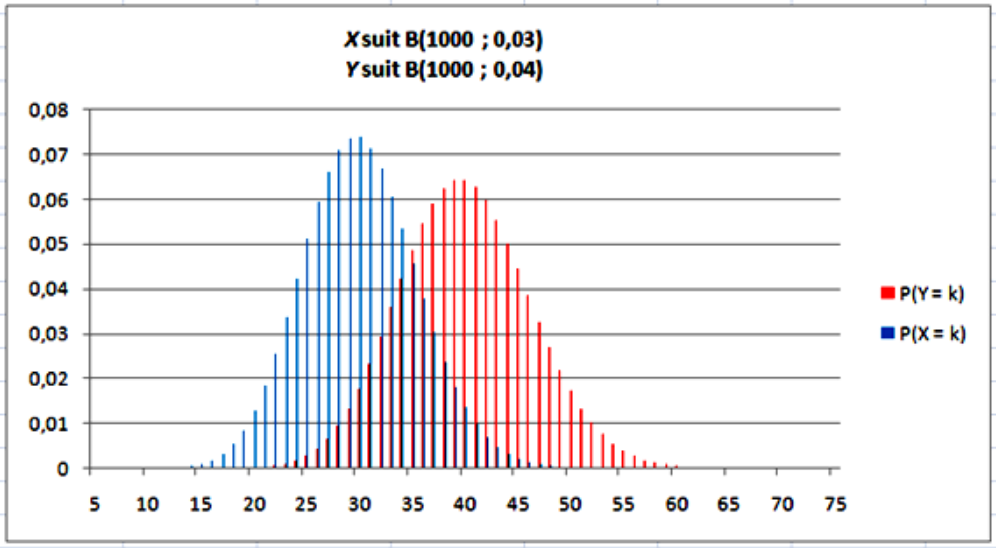
D'après une présentation d'Alain VIVIER, ingénieur-chercheur au C.E.A., chargé d'enseignement à l'INSTN.

A996 : =ENT(ALEA()+0,03)

	A	B	C	D	E	F	G	H
993	0	0	0	0	0	0	0	1
994	0	0	0	0	0	0	0	0
995	0	0	0	0	0	0	0	0
996	0	0	0	0	0	0	0	0
997	0	0	0	0	0	0	0	0
998	0	0	0	0	0	0	0	0
999	0	0	0	0	0	0	0	0
1000	0	0	0	0	1	0	0	0
1001	0	0	0	0	0	0	0	0
1002	35	31	27	36	32	27	37	45
1003								
1004	Moyenne	29,85						
1005	Pourcentage des comptages inférieurs ou égaux au seuil 39 :					97	%	
1006								

	B	C	D	E	F	G
1	uit B(1000 ; 0,03)			Y suit B(1000 ; 0,04)		
2	P(X = k)	P(X <= k)		k	P(Y = k)	P(Y <= k)
3	5,912E-14	5,912E-14		0	1,86738E-18	1,8674E-18
4	1,8285E-12	1,88757E-12		1	7,78076E-17	7,9675E-17
5	2,8247E-11	3,01343E-11		2	1,61937E-15	1,699E-15
6	2,9062E-10	3,20756E-10		3	2,24463E-14	2,4145E-14
7	2,2403E-09	2,56109E-09		4	2,33114E-13	2,5726E-13
8	1,3802E-08	1,63634E-08		5	1,93484E-12	2,1921E-12
9	7,079E-08	8,71535E-08		6	1,33692E-11	1,5561E-11
10	3,1089E-07	3,98046E-07		7	7,91013E-11	9,4663E-11
11	1,1935E-06	1,59154E-06		8	4,09102E-10	5,0376E-10
12	4,0685E-06	5,66007E-06		9	1,87884E-09	2,3826E-09
13	1,247E-05	1,81299E-05		10	7,75804E-09	1,0141E-08
14	3,471E-05	5,28398E-05		11	2,90927E-08	3,9233E-08
15	8,8474E-05	0,000141314		12	9,9905E-08	1,3914E-07
16	0,00020796	0,000349275		13	3,16366E-07	4,555E-07
17	0,00045344	0,000802714		14	9,29325E-07	1,3848E-06
18	0,00092184	0,001724552		15	2,54532E-06	3,9301E-06
19	0,00175517	0,003479726		16	6,529E-06	1,0459E-05
20	0,00314207	0,006621797		17	1,57464E-05	2,6206E-05
21	0,00530697	0,011928766		18	3,58304E-05	6,2036E-05
22	0,00848309	0,020411853		19	7,71611E-05	0,0001392
23	0,01286893	0,033280784		20	0,000157698	0,00029689
24	0,01857371	0,051854499		21	0,000306635	0,00060353
25	0,02556279	0,077417291		22	0,000568552	0,00117208
26	0,03361777	0,111035056		23	0,001007326	0,00217941
27	0,04232546	0,153360516		24	0,001708607	0,00388801
28	0,05110472	0,204465235		25	0,002779334	0,00666735
29	0,05927094	0,263736173		26	0,004342709	0,01101006
30	0,06612817	0,329864344		27	0,006527467	0,01753753
31	0,07107074	0,400935084		28	0,009451229	0,02698875
32	0,07367305	0,47460813		29	0,01319913	0,04018788





4 - La place de l’évaluation

Les compétences du « socle commun » et les enquêtes internationales

- La « culture mathématique » au sens PISA.
- Evaluation du sens critique, de la capacité à prendre des initiatives.

La prise en compte des compétences développées grâce à l'outil informatique

- Travaux pratiques évalués.
- Baccalauréat professionnel en contrôle en cours de formation comportant une partie informatique.
- Epreuve pratique au baccalauréat général.
- Vers une meilleure préparation aux études supérieures.

5 - Des résistances… et des perspectives

Des résistances

- La question de la formation des enseignants.
- L'image de la statistique chez les professeurs de mathématiques.
- La difficulté de l'interdisciplinarité.
- Le « devoir surveillé » classique.
- La difficile évolution du baccalauréat.

Des perspectives

- Une évolution « continue » des programmes (vers la loi normale et son utilisation en inférence en terminale ?).
- Une évolution de la forme des examens déjà engagée dans la filière professionnelle.
- Un équipement informatique et une formation des enseignants devenant performants en lycée.